

SELF ORGANIZING MAP (SOM) CLUSTERING UNTUK ANALISIS POTENSI OBYEK DAYA TARIK WISATA (ODTW) DI JAWA BARAT

Ahmad Fauzi¹, Eko Harli², Tria Hadi Kusmanto³

^{1,2,3}Universitas Indraprasta PGRI Jakarta

Jl. Raya Tengah No.80, Gedong, Pasar Rebo, Jakarta Timur 13760

¹ahmadfauzi.udzi@gmail.com, ²ekoharli@gmail.com, ³triahadi226@gmail.com

ABSTRAK

Sektor pariwisata di Jawa Barat memiliki potensi besar dengan keberagaman objek daya tarik wisata (ODTW) yang meliputi wisata alam, budaya, buatan, dan minat khusus. Namun, pengembangan pariwisata belum merata karena adanya ketimpangan dalam pengelolaan dan promosi antar daerah, sehingga diperlukan analisis berbasis data untuk mengidentifikasi pola dan mengelompokkan daerah dengan karakteristik serupa. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Jawa Barat berdasarkan jenis dan jumlah potensi ODTW menggunakan algoritma *Self-Organizing Map* (SOM), yang dipilih karena kemampuannya memetakan data berdimensi tinggi ke dalam peta dua dimensi yang intuitif. Data jumlah ODTW dikumpulkan, dipreproses, dan dianalisis menggunakan SOM, menghasilkan tiga kelompok utama dengan kategori tinggi, sedang, dan rendah. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi pengambil kebijakan dalam merancang strategi pengembangan pariwisata yang lebih efektif dan berkelanjutan, sehingga mampu mengoptimalkan kontribusi sektor pariwisata terhadap perekonomian daerah.

Kata Kunci: som, wisata, klaster, odtw.

ABSTRACT

The tourism sector in West Java has significant potential due to the diversity of tourist attractions, including natural, cultural, artificial, and special interest tourism. However, tourism development remains uneven due to disparities in management and promotion across regions. Therefore, data-driven analysis is needed to identify patterns and cluster areas with similar characteristics. This study aims to classify districts/cities in West Java based on the types and numbers of the diversity of tourist attractions using the Self-Organizing Map (SOM) algorithm, which was selected for its ability to map high-dimensional data into an intuitive two-dimensional representation. Data on the number of the diversity of tourist attractions were collected, pre-processed, and analyzed using SOM, resulting in three main clusters categorized as high, medium, and low. The findings of this study are expected to provide insights for policymakers in designing more effective and sustainable tourism development strategies, thereby optimizing the tourism sector's contribution to regional economic growth.

Keywords: som, tourism, clustering, odtw.

PENDAHULUAN

Indonesia merupakan negara kepulauan dengan kekayaan budaya dan keindahan alam yang luar biasa, menjadikannya salah satu destinasi pariwisata terkemuka di dunia. Sektor pariwisata berperan penting dalam perekonomian nasional, menyumbang pendapatan negara, menciptakan lapangan kerja, serta mendorong pembangunan daerah. Salah satu provinsi yang memiliki potensi besar di sektor ini adalah Jawa Barat, dengan keanekaragaman objek daya tarik wisata (ODTW) yang tersebar di berbagai kabupaten dan kota. Jawa Barat menawarkan wisata alam, budaya, wisata buatan, dan wisata minat khusus yang mampu menarik wisatawan domestik maupun mancanegara.

Namun, meskipun memiliki potensi yang besar, pengembangan sektor pariwisata di Jawa Barat menghadapi berbagai tantangan. Ketimpangan pengelolaan dan promosi antar daerah, kurangnya pemahaman terhadap keunggulan kompetitif masing-masing wilayah, serta belum adanya strategi berbasis data yang terintegrasi menjadi faktor yang menghambat optimalisasi sektor ini. Beberapa daerah dengan potensi besar belum dikembangkan secara maksimal, sementara daerah lain menghadapi stagnasi akibat promosi yang kurang efektif. Tanpa pemetaan dan pengelompokan yang jelas, pengalokasian sumber daya dan perumusan strategi menjadi kurang tepat sasaran.

Pendekatan analisis data melalui teknik clustering dapat menjadi solusi untuk mengatasi permasalahan tersebut. Dengan mengelompokkan daerah berdasarkan kesamaan potensi dan karakteristik ODTW, pemerintah daerah dapat merancang strategi yang lebih terarah dan efektif. Clustering memungkinkan identifikasi pola dan potensi tersembunyi yang sulit terlihat melalui analisis konvensional. Studi terdahulu menunjukkan bahwa penerapan clustering, seperti K-Means, telah berhasil digunakan untuk mengelompokkan destinasi wisata berdasarkan kinerja dan faktor ekonomi (Assaf & Josiassen, 2011).

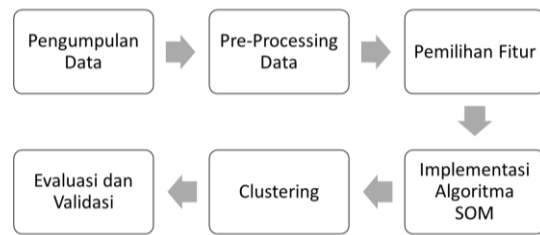
(Habiballah et al., 2024) telah melakukan penelitian terkait penerapan algoritma K-Means dalam mengelompokkan ODTW di Jawa Barat, menghasilkan tiga *cluster* utama: kategori tinggi, sedang, dan rendah. Namun, algoritma K-Means memiliki keterbatasan dalam menangkap pola *non-linear* dan hubungan kompleks antar variabel, sehingga hasil *clustering* terkadang kurang akurat untuk data dengan dimensi dan variasi tinggi.

Untuk menjawab tantangan ini, penelitian ini mengusulkan penggunaan algoritma *Self-Organizing Map* (SOM) dalam mengelompokkan kabupaten/kota di Jawa Barat berdasarkan jenis dan jumlah ODTW. SOM dipilih karena memiliki keunggulan dalam memetakan data berdimensi tinggi secara *non-linear* dan menghasilkan representasi visual yang lebih intuitif (Fawaz et al., 2022). Dengan menggunakan data jumlah ODTW yang dikategorikan ke dalam wisata alam, budaya, buatan, dan minat khusus, penelitian ini bertujuan untuk memberikan pemetaan yang lebih akurat dan informatif.

Penelitian ini bertujuan memberikan rekomendasi berbasis data bagi para pengambil kebijakan dalam perencanaan dan pengembangan sektor pariwisata Jawa Barat. Hasil penelitian ini diharapkan mampu mengidentifikasi daerah-daerah dengan potensi tinggi yang memerlukan prioritas pengembangan, serta memperkuat promosi daerah dengan kategori sedang dan rendah. Dengan demikian, penelitian ini diharapkan dapat mendukung optimalisasi sektor pariwisata, meningkatkan daya saing daerah,

dan mendorong pertumbuhan ekonomi yang berkelanjutan di Jawa Barat.

METODE PENELITIAN



Gambar 1. Tahapan Penelitian

Penelitian dimulai dengan pengumpulan data ODTW yang mencakup wisata alam, budaya, buatan, dan minat khusus. Data yang dikumpulkan kemudian diproses melalui tahap pembersihan, normalisasi, dan pemilihan fitur relevan untuk memastikan kualitas data yang optimal. Setelah itu, algoritma SOM diterapkan untuk memetakan data berdimensi tinggi ke dalam peta dua dimensi, memungkinkan visualisasi pola dan pengelompokan daerah. Hasil pemetaan ini menghasilkan tiga *cluster* utama yang menunjukkan kategori tinggi, sedang, dan rendah. Tahap akhir melibatkan evaluasi dan validasi untuk memastikan bahwa hasil clustering dapat memberikan wawasan yang akurat dan berguna bagi pengambil kebijakan dalam pengembangan pariwisata yang lebih efektif dan terarah.

Teknik Pengumpulan Data

Penelitian ini menggunakan 2 jenis data, yaitu data primer dan data sekunder. Data primer merupakan dataset yang didapat langsung dari Open Data Jabar, *dataset* tersebut memiliki 807 data dengan periode dari tahun 2014-2023. Sedangkan data sekunder didapat dari artikel-artikel dan buku-buku teoritis yang terkait

Teknik Analisis Data Penelitian

pengolahan data dilakukan dengan Teknik analisis data kualitatif dimana terdiri dari 3 proses yaitu 1) Reduksi data, 2) Triangulasi Data, 3) Penarikan Kesimpulan (Fauzi & Harli, 2017)

Algoritma SOM

Self-Organizing Mapping (SOM) adalah salah satu jenis jaringan saraf tiruan yang diperkenalkan oleh Teuvo Kohonen pada tahun 1982. SOM bekerja dengan memetakan

data berdimensi tinggi ke dalam peta dua dimensi yang memudahkan visualisasi dan interpretasi kluster (biasanya berupa *grid* 2 dimensi) (Kusumadewi & Purnomo, 2004). SOM menggunakan pembelajaran tanpa pengawasan (*unsupervised learning*) (Frigui, 2008), di mana algoritma belajar untuk mengelompokkan data berdasarkan kemiripan fitur-fitur data tersebut. SOM telah banyak digunakan dalam berbagai aplikasi, seperti pengelompokan gen, analisis pasar, dan pengelompokan data spasial.

Langkah-langkah algoritma SOM dapat dijabarkan sebagai berikut (Klir & Yuan, 2015):

1. Inisialisasi vektor input $x_1, x_2, x_3, \dots, x_n$
2. Inisialisasi neuron output sebanyak $y_1, y_2, y_3, \dots, y_n$
3. Menentukan *weight* (bobot) neuron output dengan nilai antara x_{min} dan x_{max}
4. Mengulangi langkah 5 sampai 8 hingga tidak ada update *weight* (bobot) atau telah mencapai kondisi stop (*error* terkecil).
5. Pemilihan acak salah satu data dari vektor input sebagai data training.
6. Mencari jarak terdekat dari masing-masing neuron output ke data input menggunakan rumus *euclidian distance*.

$$jarak = \sqrt{\sum_{i=1}^{10} (data[i] - bobot[i])^2}$$

Dari seluruh bobot (Di) dicari yang paling kecil jaraknya, indeks dari bobot (Di) ini disebut *winning neuron*.

7. Untuk setiap bobot w_{ij} diperbaharui bobot tetangga menggunakan rumus dengan persamaan sebagai berikut :

$$w_{ij}(t + 1) = w_{ij}(t) + a(t)[x_i - w_{ij}(t)]$$

8. Mengupdate bobot bias (*error*).
9. Simpan bobot yang telah konvergen.

HASIL DAN PEMBAHASAN

Analisis Data

Data yang digunakan merupakan data potensi objek daya tari wisata di Jawa Barat yang didapat dari situs Open Data Jabar dan merupakan open access data. Data yang tersedia merupakan data dari tahun 2014 s.d 2023, data tersebut terdiri dari 9 atribut yaitu:

1. Id
2. Kode_propinsi
3. Nama_propinsi

4. Kode_Kabupaten_Kota
5. Nama_Kabupaten_Kota
6. Jenis_Odtw
7. Jumlah_Odtw
8. Satuan
9. Tahun

Data Yang Digunakan Adalah Data Rekaman Tahun Terbaru Yaitu 2023.

Pre-Processing Data

Proses Ini Menghasilkan Data Bersih, Dimana Data-Data Yang Tidak Diperlukan Pada Penelitian Ini Dihilangkan, Selain Itu Juga Atribut-Atribut Yang Tidak Diperlukan Juga Dihilangkan Sehingga Data Yang Tersedia Menjadi Bersih. Tabel 1 Menunjukkan Sebagian Data Hasil *Pre-Processing*.

Tabel 1. Hasil Pre Processing Data

Kab/Kota	Jenis ODTW	Jumlah ODTW
Kab. Bogor	Alam	122
Kab. Bogor	Budaya	35
Kab. Bogor	Buatan	54

Transformation

Tahapan Ini Untuk Mengubah Data Mentah Menjadi Bentuk Yang Lebih Sesuai Untuk Proses Data Mining. Tabel 2 Menunjukkan Hasil Proses Dari Transformasi Data. Dimana Atribut Jenis_Odtw Dipisahkan Dalam 3 Atribut Yaitu Alam, Budaya Dan Buatan.

Tabel 2. Hasil Transformation Data

Kab/Kota	Alam	Budaya	Buatan
Kab. Bogor	122	35	54
Kab. Sukabumi	115	10	5
Kab. Cianjur	103	21	32
Kab. Bandung	77	175	61
Kab. Garut	115	40	71
Kab. Tasikmalaya	110	2	5
Kab. Ciamis	16	10	24
Kab. Kuningan	117	14	29
Kab. Cirebon	8	6	4
Kab. Majalengka	118	39	47
Kab. Sumedang	36	11	60
Kab. Indramayu	13	9	6
Kab. Subang	68	1	14
Kab. Purwakarta	26	10	27
Kab. Karawang	50	39	38
Kab. Bekasi	10	2	25
Kab. Bandung Barat	74	19	70
Kab. Pangandaran	233	103	53
Kota Bogor	9	20	9
Kota Sukabumi	3	3	11
Kota Bandung	10	10	48
Kota Cirebon	1	14	2
Kota Bekasi	3	18	39
Kota Depok	22	48	20
Kota Cimahi	3	2	4
Kota Tasikmalaya	6	26	14
Kota Banjar	13	3	11

Implementasi SOM

Penelitian ini menggunakan Bahasa Pemrograman R dalam melakukan implementasi SOM, dengan pustaka paket yang digunakan untuk penelitian ini adalah paket factoextra, kohonen dan ggplot2.

Berikut adalah hasil dari setiap langkah-langkah implementasi:

1. Ringkasan Data

Langkah ini untuk mengetahui secara deskriptif bentuk data yang digunakan. Gambar 1 menunjukkan hasil *summary data*

KOTA	ALAM	BUDAYA	BUATAN
Length:27	Min. : 1.00	Min. : 1.00	Min. : 2.0
Class :character	1st Qu.: 9.50	1st Qu.: 7.50	1st Qu.:10.0
Mode :character	Median : 26.00	Median : 14.00	Median :25.0
	Mean : 54.85	Mean : 25.56	Mean :29.0
	3rd Qu.:106.50	3rd Qu.: 30.50	3rd Qu.:47.5
	Max. :233.00	Max. :175.00	Max. :71.0

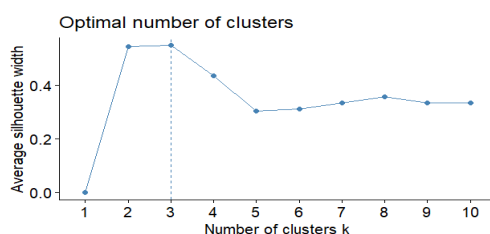
Gambar 2. Hasil Summary Data

Dapat dilihat bahwa total terdapat 27 kab/kota pada Propinsi Jawa Barat dengan nilai potensi objek daya tarik wisata alam terbanyak adalah sebanyak 233 lokasi, untuk wisata budaya terbanyak ada 175 lokasi dan wisata buatan terbanyak ada 71 lokasi

2. Menentukan Jumlah Cluster

Tahapan selanjutnya adalah menentukan jumlah *cluster* optimal dari data. Metode yang digunakan penelitian ini untuk menentukan jumlah *cluster* optimal adalah dengan menggunakan nilai *Silhouette Coefficient*. *Silhouette Coefficient* digunakan untuk melihat kualitas dan kekuatan *cluster*, seberapa baik atau buruknya suatu obyek ditempatkan dalam suatu *cluster*. Metode ini merupakan gabungan dari metode separasi dan kohesi (Ayu et al., 2019; Kodinariya & Makwana, 2013).

Gambar 2 menunjukkan bahwa jumlah *cluster* terbaik yaitu mengacu pada angka 3, walaupun terlihat nilainya juga berbeda tipis dengan jumlah *cluster* 2



Gambar 3. Jumlah Cluster

3. Clustering SOM

Setelah didapat jumlah *cluster*, kemudian dilakukan pemodelan terhadap algoritma SOM. Penelitian ini menggunakan *grid* som 5 x 5 dengan topologi heksagonal. Dalam topologi ini setiap *grid* SOM memiliki 6 tetangga terdekat, kecuali *node-node* tepi *grid*. Keuntungan menggunakan topologi ini adalah interaksi yang lebih halus dan visualisasi yang lebih efektif (Vesanto & Alhoniemi, 2000).

Dari hasil training model SOM selanjutnya dibuat *hierarchical clustering*, Gambar 3 menunjukkan hasil pengelompokan dengan *hierarchical clustering*.

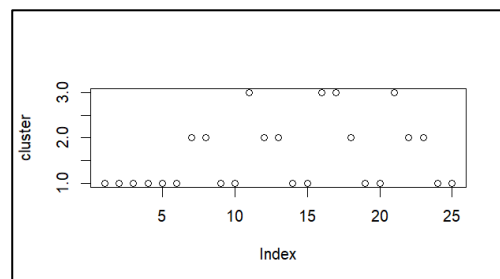
```
Call:
hclust(d = dist(som.pw$codes[[1]]))

Cluster method : complete
Distance       : euclidean
Number of objects: 25
```

Gambar 4. Hasil Pengelompokkan Hierarchical Clustering

Dari hasil tersebut didapat bahwa metode yang digunakan adalah complete dengan distance yaitu euclidean dan objek sebanyak 25 lingkaran sesuai ukuran *grid* 5x5.

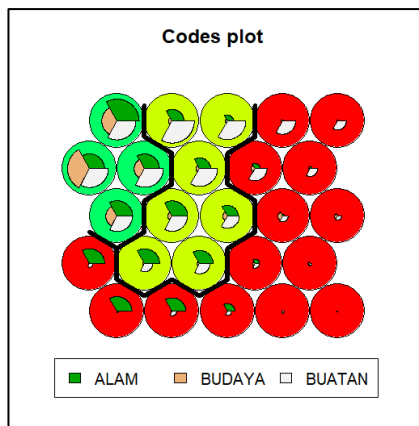
Dari hasil pengelompokan tersebut kemudian dilakukan *cutree* untuk memotong pohon hierarki pada tingkat tertentu untuk mendapatkan *cluster*.



Gambar 5. Hasil Cutree

Gambar 4. menunjukkan hasil output di atas diketahui bahwa *cluster* pertama terdiri dari 14 lingkaran, *cluster* kedua terdiri dari 7 lingkaran dan *cluster* ketiga terdiri dari 4 lingkaran. Sehingga sesuai dengan dimensi 5x5 yaitu 25 lingkaran.

Selanjutnya dilakukan visualisasi data hasil pelatihan SOM dengan menggunakan diagram Kipas seperti yang terlihat pada gambar 5 dibawah ini.



Gambar 6. Diagram Kipas Hasil Implementasi SOM

Gambar diatas merupakan hasil clustering wilayah kabupaten/kota di Provinsi Jawa Barat tahun 2023 berdasarkan faktor-faktor yang ada pada data ODTW setiap daerah menggunakan algoritma SOM. Didapatkan lingkaran berwarna merah merupakan *cluster* 1 yang beranggotakan 18 kabupaten/kota, lingkaran berwarna kuning merupakan *cluster* 2 yang beranggotakan 6 kabupaten/kota, sedangkan lingkaran hijau merupakan cluster 1 yang beranggotakan 2 kabupaten/kota.

4. Hasil Pemodelan SOM

Dari hasil clustering, didapat bahwa potensi objek wisata di Propinsi Jawa Barat dibagi kedalam 3 cluster. Dimana cluster 1 merupakan wilayah dengan potensi objek wisata kategori rendah sebanyak 18 wilayah, cluster 2 yang merupakan kategori sedang berjumlah 7 wilayah dan cluster 3 dengan kategori tinggi berjumlah 2 wilayah. Tabel 3 menunjukkan hasil dari pemodelan SOM

Tabel 3. Hasil Pemodelan SOM

Kab/Kota	Cluster	Kategori
Kab. Bogor	2	Sedang
Kab. Sukabumi	1	Rendah
Kab. Cianjur	2	Sedang
Kab. Bandung	3	Tinggi
Kab. Garut	2	Sedang
Kab. Tasikmalaya	1	Rendah
Kab. Ciamis	1	Rendah
Kab. Kuningan	2	Sedang
Kab. Cirebon	1	Rendah
Kab. Majalengka	2	Sedang
Kab. Sumedang	2	Sedang
Kab. Indramayu	1	Rendah
Kab. Subang	1	Rendah
Kab. Purwakarta	1	Rendah

Kab. Karawang	1	Rendah
Kab. Bekasi	1	Rendah
Kab. Bandung Barat	2	Sedang
Kab. Pangandaran	3	Tinggi
Kota Bogor	1	Rendah
Kota Sukabumi	1	Rendah
Kota Bandung	1	Rendah
Kota Cirebon	1	Rendah
Kota Bekasi	1	Rendah
Kota Depok	1	Rendah
Kota Cimahi	1	Rendah
Kota Tasikmalaya	1	Rendah
Kota Banjar	1	Rendah

Pembahasan

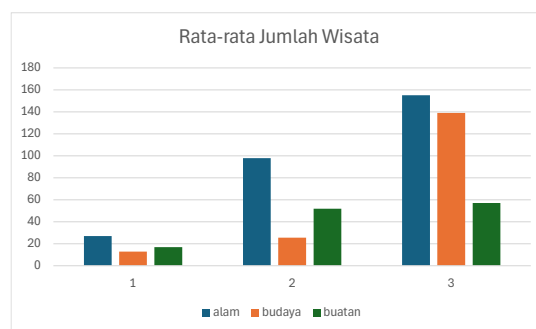
Hasil *clustering* menunjukkan 3 buah *cluster* yang didapat dari implementasi algoritma SOM pada data ODTW di Propinsi Jawa Barat. Hasil ini tidak berbeda jauh dari hasil penelitian sebelumnya (Habiballah et al., 2024).

Namun terlihat bahwa pada tahun 2023 ini ada penambahan 1 wilayah yang masuk kedalam *cluster* kategori tinggi yaitu Kab. Pangandaran. Merujuk pada data tahun sebelumnya Kab. Pangandaran mengalami kenaikan yang baik dalam hal peningkatan objek wisata. Tabel 4 menunjukkan kenaikan jenis wisata di Kab. Pangandaran.

Tabel 4. Data Kenaikan Jenis Wisata di Kab. Pangandaran

Tahun	Jenis ODTW	Jumlah Lokasi
2022	Alam	182
2022	Budaya	76
2022	Buatan	28
2023	Alam	233
2023	Budaya	103
2023	Buatan	53

Karakteristik dari setiap *cluster* pun terlihat cukup baik. Gambar menampilkan grafik rata-rata jumlah wisata pada setiap *cluster*



Gambar 7. Grafik rata-rata jumlah wisata setiap cluster

Dari gambar 6 terlihat bahwa *cluster* tinggi memiliki jumlah objek wisata yang tinggi pula. Hasil tersebut menggambarkan bahwa masih banyak Kota/Kabupaten di Jawa Barat yang belum meningkatkan sektor wisata pada wilayahnya

SIMPULAN DAN SARAN

Berdasarkan analisis pengelompokan wilayah kabupaten dan kota di Jawa Barat menggunakan algoritma SOM, ditemukan 3 *cluster* berdasarkan jenis dan jumlah potensi ODTW, yakni 2 wilayah dengan *cluster* tinggi, 6 wilayah *cluster* sedang, dan 18 wilayah *cluster* rendah. Hasil ini sebanding dengan penelitian sebelumnya yang menggunakan algoritma K-Means pada data tahun 2022, menunjukkan bahwa perkembangan objek wisata di Jawa Barat dalam setahun terakhir relatif stagnan. Oleh karena itu, disarankan agar pihak terkait lebih fokus pada pengembangan dan promosi potensi ODTW di wilayah-wilayah dengan *cluster* rendah dan sedang, serta memperhatikan distribusi objek wisata untuk merata di seluruh wilayah.

DAFTAR PUSTAKA

- Assaf, A. G., & Josiassen, A. (2011). Identifying and Ranking the Determinants of Tourism Performance: A Global Investigation. *Journal of Travel Research*, 51(4), 388–399. <https://doi.org/10.1177/0047287511426337>
- Ayu, D., Cahya, I., Ayu, D., & Pramita, K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. 9(3).
- Fauzi, A., & Harli, E. (2017). Rancang Bangun Penerapan Teknologi SMS Gateway Terintegrasi Mesin Biometrik pada Sistem Informasi Akademik. *Journal of Information Engineering and Educational Technology*, 1(2), 60. <https://doi.org/10.26740/jieet.v1n2.p60-64>
- Fawaz, Fitriasari, N. S., & Rosalia, A. A. (2022). Perbandingan Algoritma Self Organizing Map dan Fuzzy C-Means dalam clustering hasil produksi ikan PPN Karangantu. *EXPLORE (Jurnal Sistem Informasi Dan Telematika)*,

13(2), 102–109.

- Frigui, H. (2008). Clustering: Algorithms and Applications. *2008 First Workshops on Image Processing Theory, Tools and Applications*, 1–11. <https://doi.org/10.1109/IPTA.2008.4743793>
- Habiballah, H., Faqih, A., & Supriati, T. (2024). Implementasi Algoritma K-Means Dalam Mengelompokan Kabupaten/Kota Di Jawa Barat Berdasarkan Jenis Dan Jumlah Potensi Objek Daya Tarik Wisata. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 12(2), 1482–1490.
- Klir, G. J., & Yuan, B. (2015). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. <https://api.semanticscholar.org/CorpusID:10090179>
- Kusumadewi, S., & Purnomo, H. (2004). *Aplikasi logika fuzzy untuk pendukung keputusan*. Graha Ilmu.
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 11(3), 586–600. <https://doi.org/10.1109/72.846731>